Rinaldi, S., Van Der Kamp, M. W., Ranaghan, K. E., Mulholland, A. J., & Colombo, G. (2018). Understanding Complex Mechanisms of Enzyme Reactivity: The Case of Limonene-1,2-Epoxide Hydrolases. *ACS Catalysis*, *8*(7), 5698-5707. https://doi.org/10.1021/acscatal.8b00863

Understanding complex mechanisms of enzyme reactivity: the case of Limonene-1,2-epoxide hydrolases.

Silvia Rinaldi,[1] Marc W. Van der Kamp,[2,3] Kara E. Ranaghan,[2] Adrian J. Mulholland,[2*] Giorgio Colombo[1,4*]

[1] Istituto di Chimica del Riconoscimento Molecolare, C.N.R., Via Mario Bianco 9, 20131 Milano, Italy

[2] Centre for Computational Chemistry, School of Chemistry, University of Bristol, Cantock's Close, Bristol, BS8 1TS, United Kingdom

[3] School of Biochemistry, University of Bristol, University Walk, Bristol, BS8 1TD, United Kingdom

[4] Dipartimento di Chimica, Università degli Studi di Pavia, Via Taramelli 12, 27100 Pavia, Italy

Authors to whom correspondence should be addressed:

*Giorgio Colombo: g.colombo@unipv.it

*Adrian J. Mulholland: Adrian.Mulholland@bristol.ac.uk

Abstract

Limonene-1,2-epoxide hydrolases (LEHs), a subset of the epoxide hydrolase family, present interesting opportunities for the mild, regio- and stereo- selective hydrolysis of epoxide substrates. However, moderate enantioselectivity for non-natural ligands, combined with narrow substrate specificity, has so far limited the use of LEHs as general biocatalytic tools. A detailed molecular understanding of the structural and dynamic determinants of activity may complement directed evolution approaches to expand the range of applicability of these enzymes. Herein, we have combined quantum mechanics/molecular mechanics (QM/MM) free energy calculations for the reaction with MD simulations of the enzyme internal dynamics, and the calculation of binding affinities (using the WaterSwap method) for various representatives of the enzyme conformational ensemble, to show that the presence of natural or non-natural substrates differentially modulates the dynamic and catalytic behavior of LEH. The cross-talk between the protein and the ligands favors the selection of specific substrate-dependent interactions in the binding site, priming reactive complexes to select different preferential reaction pathways. The knowledge gained via our combined approach provides a molecular rationale for LEH substrate preferences. The comprehensive strategy we present here is general and broadly applicable to other cases of enzyme-substrate selectivity and reactivity.

Keywords

Biocatalysis, catalytic mechanism, stereoselective epoxide hydrolysis, free energy landscape, substrate specificity, QM/MM and MD calculations, internal dynamics, binding affinity.

Introduction

Epoxide hydrolases (EHs) are a class of enzymes that catalyze the ring opening of epoxides via addition of a water molecule, resulting in the production of vicinal diols in mild aqueous environments, achieving a regio-, stereo-, and substrate-selectivity unrivaled by synthetic catalysts.[1–3] The kinetic resolution of racemic epoxide mixtures, as well as the stereoselective hydrolysis of meso-epoxides, make these biocatalysts a valuable resource for synthetic organic chemistry and biotechnological applications.[4–8] Most EHs share high sequence and structural similarity, exhibiting a classical α/β-hydrolase fold.[9,10] However, a new EH sub-family has been recently discovered, named limonene-1,2-epoxide hydrolases (LEHs) from the natural substrate of the first isolated member.[11,12] These homo-dimeric proteins show a different structure and catalytic mechanism from the α/β hydrolase fold family.[13–16] LEH features a highly curved six-stranded mixed β-sheet, with four α-helices packed onto it to create a deep cavity. Notably, this pocket is largely hydrophobic in nature, except for a cluster of polar and charged residues at its deepest point that play a crucial role in the enzyme's reaction mechanism (Figure 1a and 1b).[17]

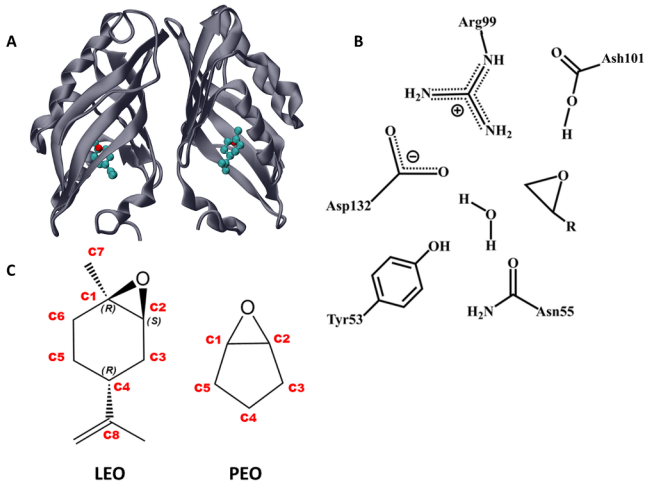

Figure 1. A) LEH dimeric structure with limonene-1,2-epoxide shown in the active site (PDB code 1NU3). B) Key residues in the LEH catalytic site. C) the natural substrate limonene epoxide (LEO) and the non-natural *meso* epoxide substrate cyclopentene oxide (PEO).

When incubated with a diastereomeric mixture of its natural substrate, LEH shows a sequential and enantio-convergent conversion of the different isomers.[18,19] Unfortunately, its use as a regio- and stereospecific biocatalyst has been limited so far by several issues, in particular: moderate to low enantioselectivity for substrates different from limonene-1,2-epoxide; narrow substrate specificity; and low thermal stability.[18,20] Several studies have been conducted in recent years to address these issues. In particular, the groups of Reetz and Janssen have exploited directed evolution-guided approaches, sometimes coupled to computational methods, to evolve and engineer LEH properties.[21–28] Although a broad set of highly enantioselective or thermostable variants has been obtained, the tradeoff between catalytic activity and structural stability,[29] combined with inadequate activity towards non-natural substrates is still hampering industrial applications of LEH.

Molecular simulation and modelling approaches have an important role to play in the development of LEH biocatalysts, because they can provide atomic-level insight into reactivity and binding, and reveal possible connections between enzyme structural and dynamic properties. Such knowledge can be used to guide the enzyme design process to engineer better properties.

The Himo group used a DFT cluster approach to build a simplified LEH active site model (*i.e.*, only some key residues around the catalytic pocket were explicitly considered) that shed light on the enzyme catalytic mechanism.[30,31] These studies demonstrated that theoretical methodologies that provide a quantum mechanical characterization of transition states can generate a quantitative description of catalytic mechanisms and can be also used as tools in the field of asymmetric biocatalysis.[31] According to Himo et al.'s results, LEH reacts by a one-step SN_2-like general acid/general base-catalyzed mechanism that significantly differs from the two-step general base-catalyzed reaction of the other EH family members; this is supported by a hybrid quantum mechanics/molecular mechanics (QM/MM) study[32]. In particular, the Asp101–Arg99–Asp132 catalytic triad appears to drive a concerted reaction in which Asp132 acts as a general base, extracting a proton from a water molecule. The resulting hydroxide then attacks the epoxide, while Asp101 protonates the oxirane ring facilitating its opening. Overall, this model involves a proton abstraction-proton donation mechanism that is well-known in enzyme catalysis and is sometimes referred to as a push-pull mechanism (Figure 1b and S1).

Although these data have elegantly clarified important features of LEH reactivity, several crucial points still remain to be addressed, such as the relevance of LEH dynamics/conformational behavior for its mechanism of substrate recognition and regioselectivity, as well as the impact of structural and conformational properties on LEH's limited adaptability to different substrates. Natural enzymes have indeed evolved to favor adaptation of their dynamic states to their catalytic/biological functions.[33,34] The processes and interactions involved in substrate recognition, ligand-dependent conformational

modifications of the active-site, the active role of water molecules, regulation at a distance via allosteric phenomena, all contribute to the complexity of the bio-catalysis problem and play a key role in fine tuning enzymatic behavior.[35–43] A comprehensive understanding of the properties of LEH therefore requires consideration of the dynamics of the whole enzymatic system.

Our purpose here is to gain mechanistic insight into LEH recognition, binding and catalysis and to investigate how these could be affected by the dynamics of the enzyme. Classical MD simulations, both in the presence of limonene-1,2-epoxide and the *meso* compound cyclopentene oxide (Figure 1c), for which LEH has very low activity,[19] were performed and the different conformations of the catalytic pocket analyzed. QM/MM methods were used to explore both the free and potential energy surfaces that govern the enzymatic reaction for the different substrates. By integrating QM/MM free energy calculations with classical MD simulation and first-principles binding affinity calculations (with Monte Carlo sampling)[44] we show it is possible to obtain a deeper understanding of the complex nature of LEH catalysis. We demonstrate that LEH reactivity is finely modulated by the nature of and interactions with substrates. The specificity of the substrate-binding guides the internal dynamics of the protein. In our model, the enzyme-substrate cross-talk stabilizes different protein states that ultimately translate into differences in both affinity and catalytic efficiency for the natural and non-natural substrates. Likewise, the substrate-dependent stabilization of preferential catalytic conformations results in different initial events of catalytic activation. The intimate relationship between protein internal dynamics, ligand binding affinity and catalytic pathways we describe for this LEH case study reflects the complexity of enzyme catalytic processes, and shows that all these aspects should be considered for a full understanding of LEH activity and selectivity that can ultimately assist the rational design of LEH variants with high efficiency for non-natural substrates.

Results and Discussion

Structural and conformational properties of LEH: the impact of substrate binding on protein dynamics. Limonene-1,2-epoxide, 1-methylcyclohexene oxide, cyclohexene oxide, and indene oxide are the only substrates accepted by the wild-type LEH.[18] Although directed evolution-guided approaches have been able to extend LEH enantioselectivity to different substrates,[21,22,24,27,45] the level of enzymatic activity still remains a significant bottleneck. In order to investigate the effect of different ligands on LEH plasticity and on the preferential stabilization of catalytically active conformational ensembles, all-atom classical MD simulations were run for three model systems: (i) the *apo* enzyme

(APO), (ii) the natural substrate limonene-1,2-epoxide (LEO), and (iii) the *meso* compound cyclopentene oxide (PEO) for which LEH shows less than 0.25% of the LEO hydrolysis rate.[18]

LEH forms a stable homodimer, and each monomer can bind a substrate molecule within its catalytic pocket:[46] therefore, a complete functional characterization should consider the whole system. Indeed, several studies have already demonstrated the crucial role of monomer-monomer interactions in stabilizing and tuning LEH dynamics and stability.[23,25,47] Therefore, unlike previous computational characterizations,[28,30–32] the whole dimer has been considered in our simulations and consequentially the different substrates have been modeled in both monomeric chains (Figure 1a). In addition to improving the reliability of the model, this allows increased sampling of the conformational landscape of the catalytic pockets. While experimental evidence[47] supports a model whereby the dimer is the active species and the starting crystal structure is doubly occupied by the natural inhibitor, it is conceivable that only one monomer could be occupied or react at any time. We thus also built a model where only one of the two monomers is occupied by the ligand. In particular, we aimed to test whether single- or double-occupancy could induce different internal dynamics (altering the conformations sampled). A comparison between the two cases is reported in the SI (see Figure S10). Considering the observed similarities in the internal dynamics patterns between the single or doubly occupied states, we proceeded with the analyses of the latter, which in any case represents the form considered more likely to be present in solution.

It is important to emphasize here that for each ligand state, we are not interested in the characterization of macroscopic conformational rearrangements such as (un)folding and (un)binding events; rather, we focus on the investigation of the enzyme internal dynamics around native, catalytically competent structures. Our aim is to identify the determinants of microscopic dynamics that may directly be related to the onset of functionally-oriented motions, differentially selected by different substrates. Note that the best substrate among the 4 stereoisomers of limonene-1,3-epoxide[18] (1R,2S,4R-limonene-1,2-epoxide) was considered in our calculations.

Whilst LEO stably occupies the LEH binding site (100% of the simulated time), the smaller PEO ligand shows significant mobility within the active site, retaining the original binding pose less than 50% of the simulated time (Figure S3). Therefore, in order to evaluate the effects of PEO on LEH structures and dynamics (and obtain an equilibrated starting structure for free energy calculations), we restrained PEO and the catalytic water to be within hydrogen-bond distance from Asp101 and Asp132, respectively (see SI) in monomer B as this is an important requirement for the ring opening reaction. No restraints were placed on PEO in monomer A (PEO-A), so that the dynamics in that subunit are unbiased.

The time evolution of the Root-Mean-Square Deviation (RMSD) compared to the X-ray structure (all backbone heavy atoms excluding the highly mobile N-terminal tail, see Figure S5) shows that the protein is stable in all the model systems (avg. RMSD of 1.2 Å, 1.2 Å and 1.5 Å for LEO, PEO and, APO respectively). The conformational variability of residues in the active site (taking as reference the residues that have at least one atom within 4 Å from LEO in the X-ray structure) in the different models (RMSD calculated on all the heavy atoms) already indicates differential stabilizing effects determined by the ligands. The presence of LEO traps the catalytic pocket in a unique structural ensemble, compared to the variable behavior of the APO state (Figure S6, avg. RMSD of 1.0 Å and 2.2 Å for the LEO and APO forms, respectively). The presence of PEO within the LEH pocket has a stabilizing effect compared to the APO form. However, the catalytic site shows greater conformational plasticity than with the natural substrate LEO, with average RMSD of 1.4 Å (PEO-A). Note that restraining PEO in the active site significantly narrows the structural ensemble populated in the MD simulation (avg. RMSD of 1.0 Å). Interestingly, the water molecule responsible for the nucleophilic attack (resolved in the crystallographic structure) also shows ligand-dependent dynamics. The presence of PEO increases the mobility of the pocket, and as a consequence, water molecules can undergo significant reorganization in the binding site. The nucleophilic water molecule in PEO-A is within 4 Å of both the Asp132-oxygen and one of epoxide-carbons only in 67% of structures, compared to 96% and 90 % for the LEO and PEO-B models, respectively. Very recent work by the Reetz group,[28] published while this paper was being submitted, describes similar behavior. Indeed, by combining crystallographic structural investigation with MD studies, those authors showed significant flexibility of the LEH binding pocket in the presence of PEO, which prevents that correct alignment for the catalysis between the substrate and the enzyme. The differences in structural mobility between LEO and PEO therefore offer a first glimpse of the conformational determinants of differential enzyme activity. The larger steric bulk of LEO compared to PEO rigidifies the binding pocket, favoring the correct positioning of the substrate, the nucleophilic water and the catalytic residues that are required by the concerted mechanism for epoxide opening. The representative structures (obtained by cluster analysis) of the binding pockets show that LEO binding promotes a broader network of hydrophobic interactions compared to the PEO state, particularly in the region of the LEO-isopropenyl substituent. The latter interacts with a hydrophobic cluster defined by Leu58-Leu74-Met78-Leu147, which accounts for the majority of the conformational diversity of the enzyme cavity in the presence of the different substrates (see Figure 2). This suggests a possible key role for these residues in the modulation of substrate specificity.

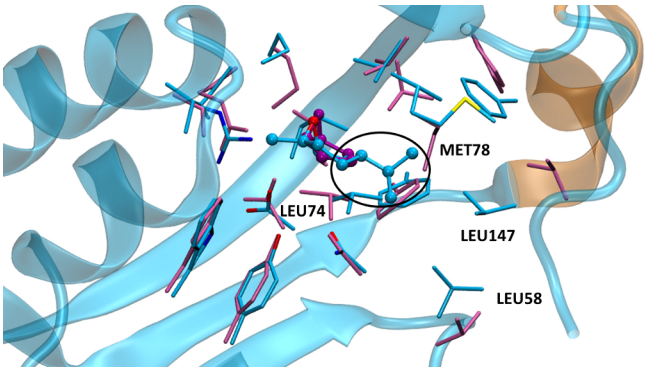

Figure 2. LEH substrate binding pocket. The residues that make up the enzyme active site that bind either to LEO or PEO are shown as sticks in blue and purple respectively. The VMD CPK representation is used for substrate molecules, and the black oval pinpoints the LEO-isopropenyl substituent. Residues belonging to the hydrophobic cluster are specifically indicated. The C-terminal α-helix is highlighted in orange.

These data are in agreement with previous studies: Reetz, Janssen and coworkers demonstrated that point mutations of residues located in this area modulate LEH activity and enantioselectivity.[21,28,45] These groups were able to obtain highly enantioselective LEH variants that catalyze the hydrolytic opening of the PEO substrate. However, any (combination of) mutations in this region did not improve on the conversion displayed by *wt* LEH. Therefore, other factors may contribute to the observed low efficiency of LEH for non-natural ligands.

To investigate LEH affinity for the different ligands, we computed absolute protein-ligand binding free energies using the *WaterSwap* method[44] on the most representative structures identified by cluster analysis (see SI). *WaterSwap* is a method, implemented in the Sire program (www.siremol.org) that allows the calculation of absolute binding affinities by exchanging the ligand with an equivalent volume of water. The free energy difference between the ligand-bound state, and the state in which the site is occupied by water, is calculated rigorously and efficiently by replica exchange thermodynamic integration methods with Monte Carlo simulations. The free energy costs for forming the binding site is not included (though could be calculated separately). This method exploits an explicit model for the treatment of water, thus including important water - ligand/protein interactions that are missing in continuum solvent approaches, but are essential in understanding LEH binding and activity.

Consistent with qualitative MD-based considerations, the binding affinity for PEO was calculated (using WaterSwap) to be significantly lower than for LEO: its binding free energy (obtained as weighted average for the different starting conformations) is half of that of the natural substrate (ΔG_{bind} = –22.8 and –11.6 kcal mol^{-1} for LEO and PEO, respectively).

The presence of a larger network of hydrophobic interactions (*vide infra*) can effectively stabilize LEO in its binding pocket and can be considered one of determinants of the significant binding energy differences between the two substrates. Next, by combining *WaterSwap* calculated affinities with structural investigations for several potentially reactive and non-reactive configurations we obtained further insights into the causes of the enzyme low efficiency in transforming the non-natural substrate. Indeed, for both active site cavities (*i.e*., the active site from one or other monomer, sampled with or without the bias of distance restraints, respectively) the best affinities computed for PEO (lowest energy) correspond to non-reactive geometries (–16.1 kcal mol^{-1} and –14.7 kcal mol^{-1} for PEO-B and PEO-A respectively). Non-reactive geometries are defined as those structures in which the distances between the epoxide atoms and reactive centers in the active site are beyond the defined threshold values for hydrogen transfer and/or nucleophilic attack (See Methods and SI). This result might capture a key factor for the observed lack of reactivity of the non-natural substrate, indicating that non-reactive conformations are favored for bound PEO; in contrast, in the LEO case, only catalytically-competent configurations of the complex are explored (In Supplementary Information, the structural details of complexes are defined, see Table S1).

The dynamic influence of a particular ligand may extend well beyond the active site and so we set out to unravel the network of enzyme residues that mechanically respond to the presence (or absence) of a particular ligand. Analysis of the matrix of distance fluctuations between residue pairs (see Methods) provides a useful method to identify coordinated motions within a protein and link the internal dynamics of LEH to the conformational effects of the ligands.[48] The study of such microscopic coordination patterns, in the ensemble of reactive structures, *i.e.,* structures in which the distances and relative orientations among reactive centers are compatible with the onset of acid-base and nucleophilic attack (see also Methods and the Reaction Mechanism paragraph), has proven to efficiently identify the substructures that sustain the onset of functionally-oriented motions.[49–51] Comparisons of the distance fluctuation matrices highlight interesting differences (Figure S7). The binding of LEO clearly alters the global internal dynamics of the enzyme, showing higher rigidity and an extended coordination with respect to the APO form. This is particularly noticeable in the terminal regions of each monomer, where low coordination regions (yellow areas in figure) corresponding to the N- and C- terminal tails that are visible in the APO matrix, disappear upon LEO binding. PEO appears to be an intermediate case, confirming that LEO has a stronger influence on protein dynamics. The rigidification of the active site upon LEO binding due to optimal hydrophobic packing diffuses throughout the whole protein structure. It is worth noting here the difference between the internal dynamics of monomer A (unconstrained) and B (constrained) in the PEO distance fluctuation matrix. The increased number of dark areas (highly coordinated regions) in the B monomer suggests

that the coordination (internal rigidity of the protein) increases when the ligand is bound stably to its pocket in a conformation compatible with and pre-organized for subsequent reactivity.

We then used distance fluctuation analysis to focus on LEH binding site residues whose coordination is most differentiated upon binding of PEO vs. LEO. The distance fluctuation difference matrix between the PEO-A and LEO states was calculated and a cutoff was applied to the resulting difference to select the most modulated regions. From these regions, we further filter out only the residues that are dynamically coordinated with the binding cavities. This analysis shows that the presence of LEO generally rigidifies the binding site dynamics, strongly influencing the Leu58-Leu74-Met78-Leu147 hydrophobic cluster (See Figure S8 and S9), in agreement with cluster analysis. Specifically, the presence of LEO reinforces the dynamical coordination between the hydrophobic cluster and the C-terminal α-helix (residues 139-145, orange in Figure 2). In the PEO model, the increased flexibility of Leu147-Met78 affects the local dynamics at the C-terminal region, which parallels the modulation of its secondary structure, shown by a drop in the α-helix propensity during the simulation for the PEO-A model (Figure S11). It has previously been suggested that the flexibility of this segment modulates the activity of the enzyme by regulating the entrance to the active site and therefore blocking/favoring substrate binding and product release.[25,46] Support for these observations comes from the work of Reetz and coworkers, who observed a link between the conformational changes in this region and perturbation of the active site.[21,27,28]
Our data suggest that the hydrophobic core near the binding site is allosterically connected to specific distal regions of the protein: perturbation of the properties of the latter then can be expected to have an impact not only their structural plasticity and stability but the substrate recognition and stereochemical outcome of reactions catalyzed by LEH.

Reaction mechanism: combining enzyme dynamic modulation and epoxide hydrolysis. Specificity and reactivity are two key concepts in enzyme catalysis. Many proteins are shown to adopt their catalytically active conformation only when bound to the substrates, but in other cases, enzymes explore catalytically-relevant conformations even in the absence of ligands.[52–56] Thus a complete understanding of a catalytic cycle should also take into account the conformational dynamics of the enzyme-substrate complex.

QM/MM methods can be used in conjunction with enhanced sampling techniques, such as umbrella sampling, to provide activation barriers (activation enthalpies and free energies) in good agreement with experiments.[52,57–64] QM/MM methods allow the electronic level description of the groups directly involved in the reaction, while the protein (and solvent) environment is modeled at a simpler, more

efficient, molecular mechanics level.[65,66] Here, we applied QM/MM MD umbrella sampling simulations and QM/MM adiabatic mapping methods to model LEH reactivity and map both potential and free energy profiles (potential of mean force, PMF). The level of QM theory is restricted to the semi-empirical level due to the high computational cost of extensive sampling. We used the self-consistent charge density functional tight binding (SCC-DFTB) method[67], an approximate density functional theory (DFT) based method. DFT[62] (B3LYP/6-31G*/ff99SB//SCC-DFTB/ff99SB) single point calculations are used to correct the potential energy surfaces at higher level of theory (see Methods). Similar approaches have been successfully used for the characterization of analogous systems.[68–70]

Hydrolysis by LEH occurs via a complex mechanism (see Figure S1): the Asp101–Arg99–Asp132 triad drives a concerted reaction involving the deprotonation of a water molecule by Asp132, the nucleophilic attack of the resulting hydroxide ion on the epoxide and protonation of the oxirane ring by the protonated Asp101 (specifically labeled Ash101). Arg99 is strongly associated through hydrogen bonds and electrostatic interactions with both Asp101 and Asp132 and even if it is not directly involved in the reaction mechanism, its mutation results in a deactivated enzyme.[17] This complex picture is completed by the proper positioning and activation of the nucleophilic water by the H-bond network formed by Asn55 and Tyr53 (and Asp132 itself)[17,30]. In our model, the side chains of the residues of the catalytic triad, the water molecule and the epoxide are included in the QM region. The opening of the epoxide can result from the attack on either of the two carbon atoms of the LEO oxirane ring (C1, C2, Figure 1c); however, experimental evidence indicates that the water molecule privileges the attack on the more substituted C1 atom.[30,32] Both reaction pathways were computed for the LEO model (C1-LEO, C2-LEO) while only one carbon attack was considered for the PEO *meso*-epoxide.

To model the epoxide ring opening reaction, umbrella sampling QM/MM MD simulations (SCC-DFTB/ff99SB) were performed along the following two reaction coordinates (RCs). The first reaction coordinate describes proton abstraction from the water molecule and the nucleophilic attack, and the second accounts for the activation of the epoxide by proton transfer and the ring opening. (Figure S3 and Methods). Using the transition state (TS) conformations located on the resulting 3D PMF surfaces as starting points, potential energy surfaces (PES) were obtained by adiabatic mapping along the reaction coordinates at the same level of theory (see Method sections). DFT QM/MM (B3LYP/6-31G*//ff99SB) single point energy calculations were used to correct the resulting 513 (27x19) points-spaced 3D PES.

For all the 3 cases studied (C1-LEO, C2-LEO and PEO), there are only two minima in both the free and potential energy surfaces (see Figure 3 and SI) corresponding to the reactants and diol product, consistent with the concerted mechanism proposed in the literature. However, the minimum energy

paths (MEPs) for each model indicate that the reaction for attack at C1 follows a slightly different catalytic cycle.

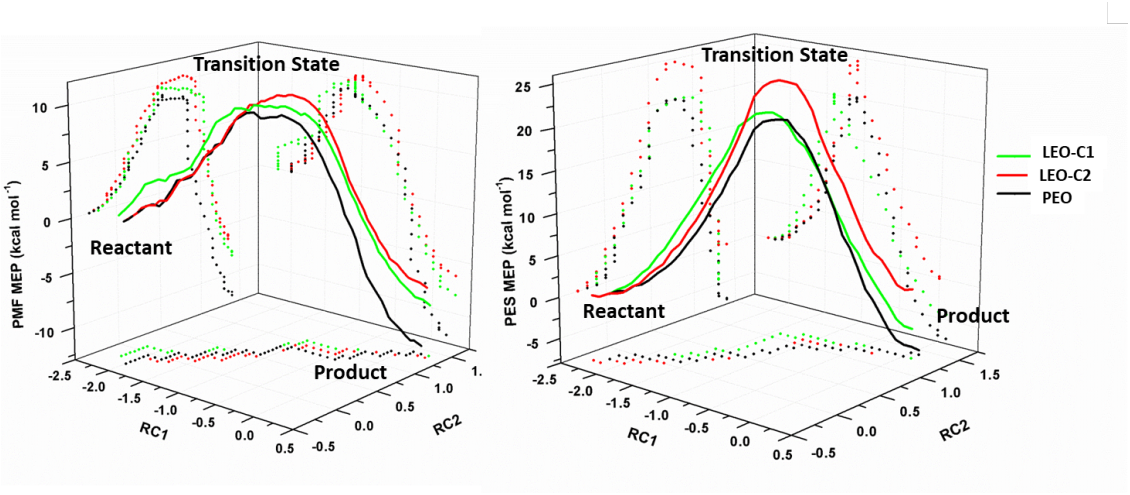

Figure 3. Left. SCC-DFTB/ff99SB Minimum Free Energy Path (MEP) as obtained by umbrella sampling simulations. Right. Minimum Potential Energy Path (B3LYP/6-31G*/ff99SB//SCC-DFTB/ff99SB) obtained by means of DFT QM/MM single-point calculations. The dots in figure report the projection on each dimension of the different MEP profiles.

Potential energy surfaces indicate (Figure 3) that the TS that drives C1-LEO attack is closer to the reactant along the first reaction coordinate and shifted towards the product along the second coordinate, compared to the C2-LEO and PEO systems (RC1= –1.4 Å/ RC2=0.8 Å in C1-LEO TS, RC1= –1.1 Å / RC2=0.7 Å in C2-LEO TS and RC1= –1.1 Å / RC2=0.7 Å in PEO TS located on DFT PESs). The resulting activation barrier ($\Delta G^{\ddagger}$) of 14.7 kcal mol^{-1} for the C1-LEO model is in reasonable agreement with the experimental values of 12.4 kcal mol^{-1}, and more importantly, reproduces the experimental trend for the relative barriers for the attack on the two epoxide carbons, being *ca*. 2.5 kcal mol^{-1} lower than the energy barrier that controls the C2-LEO attack. Surprisingly, the computed barrier for PEO opening shows an activation energy comparable to that of C1-LEO attack (14.5 kcal mol^{-1}). This outcome is in line with the recent results from the Reetz group, whereby the authors used simplified static QM models to compute the activation energies of a LEH mutant for cyclopentene and cyclohexene oxide hydrolysis.[28] They computed similar activation energies for the two systems (13.4 kcal mol^{-1} and 14.1 kcal mol^{-1} for the 5- and 6- membered ring substrates respectively, considering the attack on the most favorable carbon atom of the epoxide), while the percentage of substrate conversion still favors the hydrolysis of the cyclohexene oxide substrate. This further supports our model. Combined with previous observations, this suggests that the very low LEH activity towards PEO may not be due solely to a difference in activation energy. Other factors may come into play, including the above calculated lower binding affinity of PEO. Alternative/complementary factors include a possible slower substrate approach or product release for PEO, and the difficulty in finding

the correct positioning in the active site of the enzyme which has evolved to accept and process larger, more hydrophobic substrates. We suggest that these elements may be targeted in the design of improved LEH variants.

Of the three substrates, the reaction of PEO is the most exergonic; this can be explained by the larger amount of strain energy released upon product formation in the 5 membered-ring ligand compared to LEO; indeed, the ring strain energies per C-atom released (calculated with the MacroModel Schrodinger suite, Schrödinger Release 2018-1: MacroModel, Schrödinger, LLC, New York, NY, 2018) are *ca* 4 and 7.1 kJ mol^{-1} for LEO and PEO, respectively, in line with previous literature data.[71]

In order to rationalize the differences in the free energy surfaces, we analyzed and compared the geometry and the electronic features of the C1-LEO, C2-LEO and PEO TSs (note that TS structures derived on B3LYP/6-31G*/ff99SB//SCC-DFTB/ff99SB Potential Energy Surfaces have been used as reference for the characterizations, Figure S15). According to the previously reported model by the Himo group, epoxide opening at C1/C2 is controlled by conformational factors, in which the C1-attack results in a more stable chair-like TS with respect to the C2-attack derived twist-boat conformation.[30] The same trend is observed in our calculations, where C1-nucleophilic attack is mediated by a chair-like LEO conformation, while C2-LEO TSs structure shows a less a stable twist-boat conformation (Figure S15). The PEO TS exhibits a stable envelope-like conformation, consistent with the activation barrier values comparable to the C1-LEO pathway. Comparison of the geometrical parameters of the 3 TS structures provides further rationalization of the differences between the C1-LEO, C2-LEO and PEO energy landscapes (Figure 4 and Table S2 and S3). In agreement with the shifts in C1-LEO MEP (Figure 3), the distances between H and O atoms in Asp101 and the oxirane oxygen (Figure 4) indicate that the hydrogen transfer process (RC2) is more advanced in the TS for the C1-LEO pathway, compared to the other two models. Consequently, a significant redistribution of the electron density (charge-transfer like state) is observed (see Table S4) that leads to a rearrangement of the bond lengths pattern for the C1-LEO TS, missing in either the C2-LEO or PEO pathways (see Figure 4 and Table S2).

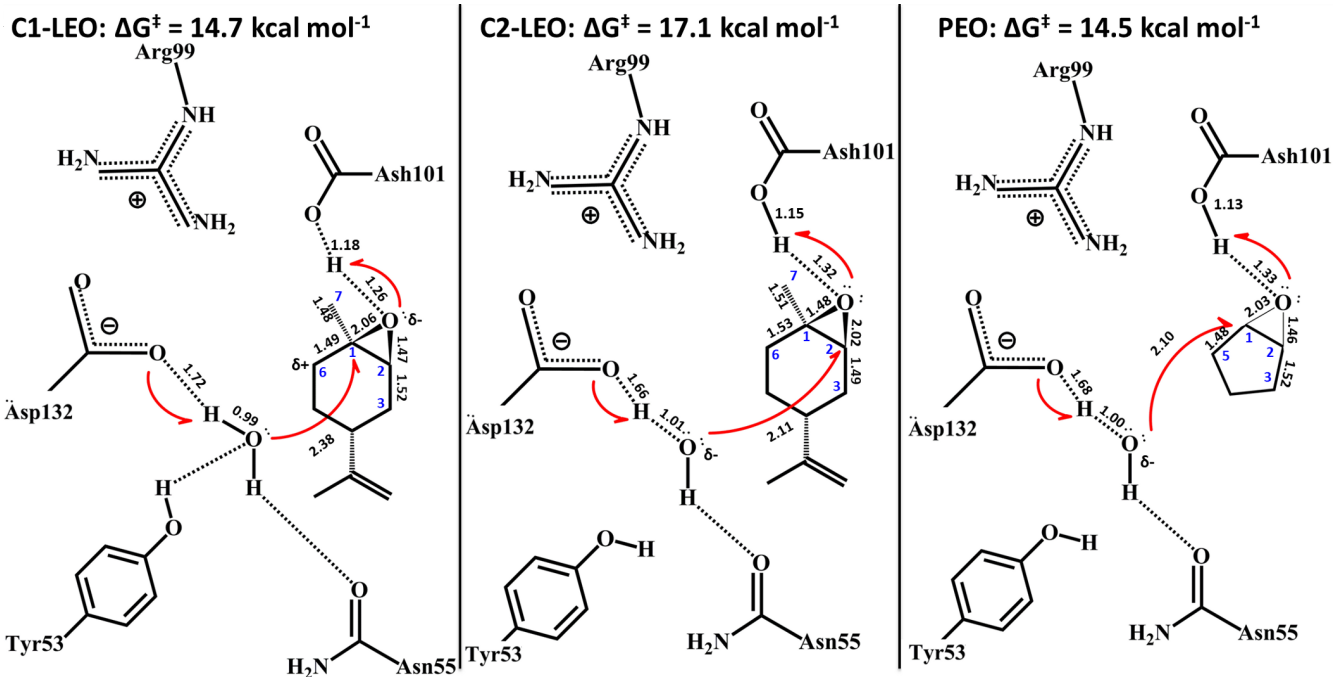

Figure 4. C1-LEO (attack on the most substituted carbon atom of limonene-1,2-epoxide), C2-LEO (attack on the less substituted carbon atom of limonene-1,2-epoxide), and PEO (*meso* compound cyclopentene oxide) TSs. Important distances of the catalytic mechanism are reported. Relevant carbon atoms of the epoxide ligands are shown in blue. Note that the values shown are taken from TS structures derived on B3LYP/6-31G*/ff99SB//SCC-DFTB/ff99SB Potential Energy Surfaces.

In particular, C1-LEO shows a stretching of the C1-O bond increasing the acid character of the oxygen atom, whose polarization (quantified here as difference between the Mulliken atomic charges of the atoms defining the covalent bond) is favored by the ability of the more substituted carbon to stabilize the developing positive charge by inductive effects (further indicated by the contraction of the C1-C7 bond, see Figure 4). The electronic redistribution also includes the contraction of the C1-C6 bond accumulating a partial positive charge on C6 (1.51 Å and 1.49 Å in C1-LEO reactant and TS respectively, see Table S2), the reduction of the C2-O distance and the consequent stretching of the adjacent C2-C3 bond by inductive effect (1.50 Å and 1.52 Å in C1-LEO reactant and TS respectively, see Table S2) complete the electron redistribution pattern in C1-LEO TS (Table S2 and Table S4).

The different reactivity of C2-LEO vs. PEO leads to different characteristics of their respective TSs. In both the C2-LEO and PEO models, protonation of the oxirane ring by Asp101 is not very advanced, but the water activation and nucleophilic attack processes are quite advanced in the TSs (Figure 4). Here, a different pattern of the H-bond network that stabilizes the nucleophilic water molecule probably increases the pK_a of water with respect to the C1-LEO attack. While the water oxygen does not act as H-bond acceptor with Tyr53 (as in C1-LEO TS), the distances between water hydrogens and their respective electronegative atoms on Asn55 (1.74 Å, 1.68 Å, 1.73 Å in C1-LEO, C2-LEO and PEO

respectively) and Asp132 are reduced (1.72 Å, 1.66 Å, 1.68 Å in C1-LEO, C2-LEO and PEO respectively, see Figure 4 and Table S3). This conformation favors the nucleophilic attack, and accordingly the water-epoxide distance in C2-LEO and PEO is significantly reduced compared to C1-LEO TS (2.38 Å, 2.11 Å, 2.10 Å in C1-LEO, C2-LEO and PEO respectively). The delocalization of the OH^- charge into the epoxide ring does not lead to the same extent of charge redistribution observed in the charge-transfer-like path for C1-LEO and the nucleophilic attack on the less substituted carbon is favored by steric factors, being this the most accessible position (and leading to an optimal nucleophilic attack angle of ca. 150°, see Table S3). In this scenario, the lack of an electron-donating group that can stabilize the positively-charged TS leads PEO to follow the C2-like reaction pathway.

Overall, the QM/MM free energy simulations show that the nature and positioning of the substrates modulates the selection of the catalytically active conformation of the enzyme binding pocket, which translates into different reaction pathways. The first event of the catalytic activation controls the regiochemistry of the epoxide opening and responds to the conformations of the binding pocket: when electronic factors become dominant, acid-driven reaction is favored, resulting in the attack on the most-substituted carbon atom (C1-LEO case). In contrast, when steric factors prevail in guiding the catalytic positioning, the attack on the less-substituted carbon atom is preferred through base-driven catalysis.

Conclusions

The simultaneous optimization of catalytic efficiency and selectivity, structural stability, and adaptability towards non-natural substrates is still a significant hurdle for the general applicability of LEH biocatalysis. Thorough understanding of the LEH functions requires the construction of a comprehensive model that can account for both dynamic and mechanistic properties and consider the long-range effects that originate from or may impact on the conformational rearrangements of the catalytic site.

Combining QM/MM free energy simulations, the characterization of the protein internal dynamics and binding affinity calculations via the first-principles *WaterSwap* method, we have explored mechanistic aspects of LEH catalysis in depth, focusing in particular on identifying correlations between enzyme function, the affinity for the ligand, function-related conformational sampling and catalysis, which may be modulated by the cross-talk between the enzyme structure and the different substrates. In addition, kinetics of substrate binding or product release can play a role.

Our results show that LEH exploits a concerted mechanism to catalyze epoxide opening; corroborating previous findings,[30,31] we have shown that LEH regioselectivity is guided by two different transition states: the preferred attack on the most substituted carbon atom is mediated by a chair-like LEO conformation, while a less stable twist-boat conformation leads the attack on the less substituted carbon atom. In addition, the (ligand dependent) organization of the active site results in a different first event of the catalytic activation, and this in turn defines the regiochemical outcome of the hydrolysis reaction. We thus suggest that LEH regioselectivity results from a combination of both conformational and electronic parameters, and the relative influence of these will depend on the specific substrate.

Our data further support the hypothesis that LEH substrate selectivity and specificity are governed by different driving forces. While the computed reaction barriers and the underlying reaction pathways for PEO and LEO hydrolysis are very similar, the lower binding affinity for the non-natural substrate substantially disfavors its reactivity.[18] The presence of the larger LEO ligand stabilizes a network of (mainly hydrophobic) interactions that favors the catalytically competent conformation. PEO binding prevents the enzyme from stably populating the active conformations: in particular, a large rearrangement is observed in the region of the hydrophobic cluster formed by Leu58-Leu74-Met78-Leu147 residues, which turns out to be an important factor for LEH specificity.[17,27,28]

Our work (which applies the *WaterSwap* method to biocatalysis for the first time) provides novel insights into the relationships between protein internal motions, molecular recognition and catalysis. The in-depth analysis of LEH catalysis sheds light on the balance between LEH selectivity and specificity for natural and non-natural substrates, and offers a basis for the rational modulation of LEH properties.

Methods

System setup and MD simulations. The starting structures for MD simulations in complex with the different ligands (LEO and PEO) were obtained from docking using the limonene-1,2-epoxide hydrolase crystallographic structure in its dimeric form in complex with the inhibitor valpromide (LEH, PDB code 1NU3, 1.75 Å resolution)[46] Docking calculations were performed using Glide,[72] part of the Schrodinger Maestro suite (Release 2013-1-9.4, Schrödinger, LLC, New York, NY, 2013) in Standard Precision mode (SP) with the OPLS-AA[73](2001) force field. The co-crystallized natural inhibitor valpromide was used to guide the docking runs. Besides providing a good starting point for the correct positioning of the ligand, this offers the opportunity to target a cavity where the conformation of protein

binding residues has already been rearranged to host a large ligand, as the comparison between the *apo* and the bound crystal structures has shown (Figure S1).[17]. The best poses of the docking runs, according to the Emodel empirical scoring function implemented in Maestro were then used as starting points for all-atom MD simulations. For all the simulations, the AMBER 12.0 package was used with the AMBER ff99SB[74] and with GAFF[75] for the ligands. Ligand parameters and AM1BCC partial charges were generated using Antechamber. The two receptor-ligand complexes and the receptor without any ligand (*apo* form) were treated using the same MD protocol (see SI). LEO and PEO MD simulations were run in 3 independent replicas for a total of 300 ns of simulated time. 100 ns of a single MD run were computed for the LEH *apo* system.

Binding affinity calculations. *WaterSwap* (replica exchange thermodynamic integration Monte Carlo) calculations[44] were performed using the *WaterSwap* executable in Sire (https://siremol.org/index.html) using the same force-field parameters and solvent model as the dynamics simulations. *WaterSwap* binding free energies were average weighted by cluster occupation and computed on the most representative structures returned by cluster analysis on each system (2 and 6 individual *WaterSwap* runs were performed for LEO and PEO respectively). See SI for further details.

QM/MM calculations. All the QM/MM MD simulations were run using the sander code available in the Amber12 package. The final structure from the MD simulations was selected as the initial geometry for the QM/MM calculations. The side-chains of the catalytic residues (namely, Asp132, Ash101, Asn55, Tyr53, Arg99), the nucleophilic water molecule and the ligand were included in the QM selections, while the rest of the protein and solvent was treated at the molecular mechanical level. Only one of the two catalytic pockets was treated by QM. Short unbiased QM/MM MD runs were performed to obtain representative structures for the subsequent umbrella sampling calculations (see SI). Two reaction coordinates (RCs) were selected in order to describe LEH catalysis by QM/MM umbrella sampling calculations (SCC-DFTB/ff99SB). The first coordinate described proton abstraction from the water molecule and the nucleophilic attack, defined as RC1= d1(Owat-Hwat) – d2(Owat-Clim). The second reaction coordinate accounted for the activation of the epoxide by proton transfer and the ring opening: RC2 = d3(Olim-Clim) – d4(Hash-Olim) (Figure S3; see SI for further details). The weighted histogram analysis method (WHAM)[75] was used to obtain the resulting PMF. Using the transition state-like conformations located on the resulting 3D PMF surface as starting points, potential energy surfaces (PESs) were obtained by adiabatic mapping along the reaction coordinates at the same level of theory (SCC-DFTB/ff99SB). Finally, *ab initio* DFT QM/MM (B3LYP/6-31G*/ff99SB//SCC-DFTB/ff99SB) single points calculations (electrostatic embedding) were run on the resulting geometries, using Orca package[76] implemented in Amber, obtaining 513 (27x19) points-spaced 3D PES (see SI).

Distance fluctuation analysis. Distance fluctuation analysis describes the dynamic coordination between any two residues.[48] It is defined as the time-dependent mean square fluctuation of the distance r_{ij} between Cα atoms of residues i and j:

Distance fluctuation$_{ij}$ = $\langle (r_{ij} - \langle r_{ij} \rangle)^2 \rangle$

Where brackets indicate the time-average over the trajectory.

ASSOCIATED CONTENT

Supporting Information

Detailed description of computational techniques, structural and dynamics characterization of MD runs, representation of energy surfaces and summary of computed energy, TS conformations and parameters, list of Cartesian coordinates for B3LYP/6-31G*/ff99SB//SCC-DFTB/ff99SB computed structures. This information is available free of charge on the ACS Publications website. PDB structures and trajectories files are available upon request.

Acknowledgements

The authors thank Dr. Jacopo Sgrignani for the helpful discussions. GC and SR acknowledge funding from "SusChemLombardia: prodotti e processi sostenibili per l'industria lombarda project, Accordo Quadro Regione Lombardia–CNR, Proj. Nr. 18096." SR acknowledges funding from AIRC through fellowship 18216. MWvdK is a BBSRC David Phillips Fellow (BB/M026280/1). We thank the Advanced Computing Research Centre of the University of Bristol for use of the computational facilities. AJM thanks EPSRC (grant number EP/M022609/1) and (with KER) BBSRC for funding (grant number BB/M000354/1).

References

(1) Weijers, C. A. G. M.; de Bont, J. A. M. Epoxide Hydrolases from Yeasts and Other Sources: Versatile Tools in Biocatalysis. *J. Mol. Catal. B Enzym.* **1999**, *6*, 199–214.

(2) Fretland, A. J.; Omiecinski, C. J. Epoxide Hydrolases: Biochemistry and Molecular Biology. *Chem. Biol. Interact.* **2000**, *129* , 41–59.

(3) Morisseau, C.; Hammock, B. D. EPOXIDE HYDROLASES: Mechanisms, Inhibitor Designs, and Biological Roles. *Annu. Rev. Pharmacol. Toxicol.* **2005**, *45*, 311–333.

(4) Kotik, M.; Archelas, A.; Wohlgemuth, R. Epoxide Hydrolases and Their Application in Organic Synthesis. *Curr. Org. Chem.* **2012**, *16*, 451–482.

(5) Widersten, M.; Gurell, A.; Lindberg, D. Structure–function Relationships of Epoxide Hydrolases and Their Potential Use in Biocatalysis. *Biochim. Biophys. Acta - Gen. Subj.* **2010**, *1800*, 316–326.

(6) Bornscheuer, U. T.; Kazlauskas, J. R. *Hydrolases in Organic Synthesis: Regio- and Stereoselective Biotransformations*; Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany **2006**.

(7) Mischitz, M.; Kroutil, W.; Wandel, U.; Faber, K. Asymmetric Microbial Hydrolysis of Epoxides. *Tetrahedron: Asymmetry* **1995**, *6*, 1261–1272.

(8) Archelas, A.; Furstoss, R. Synthetic Applications of Epoxide Hydrolases. *Curr. Opin. Chem. Biol.* **2001**, *5*, 112–119.

(9) van Loo, B.; Kingma, J.; Arand, M.; Wubbolts, M. G.; Janssen, D. B. Diversity and Biocatalytic Potential of Epoxide Hydrolases Identified by Genome Analysis. *Appl. Environ. Microbiol.* **2006**, *72*, 2905–2917.

(10) Ollis, D. L.; Cheah, E.; Cygler, M.; Dijkstra, B.; Frolow, F.; Franken, S. M.; Harel, M.; Remington, S. J.; Silman, I.; Schrag, J.; Sussman, J.; Verschueren, K.H.G.; Goldman, A. The α / β Hydrolase Fold. *Protein Eng. Des. Sel.* **1992**, *5*, 197–211.

(11) Barbirato, F.; Verdoes, J. C.; de Bont, J. A. .; van der Werf, M. J. The *Rhodococcus Erythropolis* DCL14 Limonene-1,2-Epoxide Hydrolase Gene Encodes an Enzyme Belonging to a Novel Class of Epoxide Hydrolases. *FEBS Lett.* **1998**, *438*, 293–296.

(12) van der Werf, M. J.; Swarts, H. J.; de Bont, J. A. Rhodococcus Erythropolis DCL14 Contains a Novel Degradation Pathway for Limonene. *Appl. Environ. Microbiol.* **1999**, *65*, 2092–2102.

(13) Holmquist, M. Alpha Beta-Hydrolase Fold Enzymes Structures, Functions and Mechanisms. *Curr. Protein Pept. Sci.* **2000**, *1*, 209–235.

(14) Lonsdale, R.; Hoyle, S.; Grey, D. T.; Ridder, L.; Mulholland, A. J. Determinants of Reactivity and Selectivity in Soluble Epoxide Hydrolase from Quantum Mechanics/Molecular Mechanics Modeling. *Biochemistry* **2012**, *51*, 1774–1786.

(15) Amrein, B. A.; Bauer, P.; Duarte, F.; Janfalk Carlsson, Å.; Naworyta, A.; Mowbray, S. L.; Widersten, M.; Kamerlin, S. C. L. Expanding the Catalytic Triad in Epoxide Hydrolases and Related Enzymes. *ACS Catal.* **2015**, *5*, 5702–5713.

(16) Bauer, P.; Carlsson, Å. J.; Amrein, B. A.; Dobritzsch, D.; Widersten, M.; Kamerlin, S. C. L. Conformational Diversity and Enantioconvergence in Potato Epoxide Hydrolase 1. *Org. Biomol. Chem.* **2016**, *14*, 5639–5651.

(17) Arand, M.; Hallberg, B. M.; Zou, J.; Bergfors, T.; Oesch, F.; Werf, M. J. van der; Bont, J. A. M. de; Jones, T. A.; Mowbray, S. L. Structure of Rhodococcus Erythropolis Limonene-1,2-Epoxide Hydrolase Reveals a Novel Active Site. *EMBO J.* **2003**, *22*, 2583–2592.

(18) van der Werf, M. J.; Orru, R. V. A.; Overkamp, K. M.; Swarts, H. J.; Osprian, I.; Steinreiber, A.; De Bont, J. A. M.; Faber, K. Substrate Specificity and Stereospecificity of Limonene-1,2-Epoxide Hydrolase from Rhodococcus Erythropolis DCL14; an Enzyme Showing Sequential and Enantioconvergent Substrate Conversion. *Appl. Microbiol. Biotechnol.* **1999**, *52*, 380–385.

(19) van der Werf, M. J.; Overkamp, K. M.; de Bont, J. A. M. Limonene-1,2-Epoxide Hydrolase from Rhodococcus Erythropolis DCL14 Belongs to a Novel Class of Epoxide Hydrolase. *J. Bacteriol* **1998**, *180*, 5052–5057.

(20) Ferrandi, E. E.; Sayer, C.; Isupov, M. N.; Annovazzi, C.; Marchesi, C.; Iacobone, G.; Peng, X.; Bonch-Osmolovskaya, E.; Wohlgemuth, R.; Littlechild, J. A.; Monti, D. Discovery and Characterization of Thermophilic Limonene-1,2-Epoxide Hydrolases from Hot Spring Metagenomic Libraries. *FEBS J.* **2015**, *282*, 2879–2894.

(21) Sun, Z.; Lonsdale, R.; Kong, X.-D.; Xu, J.-H.; Zhou, J.; Reetz, M. T. Reshaping an Enzyme Binding Pocket for Enhanced and Inverted Stereoselectivity: Use of Smallest Amino Acid Alphabets in Directed Evolution. *Angew. Chemie* **2015**, *127*, 12587–12592.

(22) Zheng, H.; Kahakeaw, D.; Acevedo, J. P.; Reetz, M. T. Directed Evolution of Enantioconvergency: The Case of an Epoxide Hydrolase-Catalyzed Reaction of a Racemic Epoxide. *ChemCatChem* **2010**, *2*, 958–961.

(23) Wijma, H. J.; Floor, R. J.; Jekel, P. A.; Baker, D.; Marrink, S. J.; Janssen, D. B. Computationally Designed Libraries for Rapid Enzyme Stabilization. *Protein Eng. Des. Sel.* **2014**, *27*, 49–58.

(24) Zheng, H.; Reetz, M. T. Manipulating the Stereoselectivity of Limonene Epoxide Hydrolase by Directed Evolution Based on Iterative Saturation Mutagenesis. *J. Am. Chem. Soc.* **2010**, *132* , 15744–15751.

(25) Floor, R. J.; Wijma, H. J.; Jekel, P. A.; Terwisscha van Scheltinga, A. C.; Dijkstra, B. W.; Janssen, D. B. X-Ray Crystallographic Validation of Structure Predictions Used in

Computational Design for Protein Stabilization. *Proteins Struct. Funct. Bioinforma.* **2015**, *83*, 940–951.

(26) Wijma, H. J.; Janssen, D. B. Computational Design Gains Momentum in Enzyme Catalysis Engineering. *FEBS J.* **2013**, *280*, 2948–2960.

(27) Sun, Z.; Lonsdale, R.; Wu, L.; Li, G.; Li, A.; Wang, J.; Zhou, J.; Reetz, M. T. Structure-Guided Triple-Code Saturation Mutagenesis: Efficient Tuning of the Stereoselectivity of an Epoxide Hydrolase. *ACS Catal.* **2016**, *6*, 1590–1597.

(28) Sun, Z.; Wu, L.; Bocola, M.; Chan, H. C. S.; Lonsdale, R.; Kong, X.-D.; Yuan, S.; Zhou, J.; Reetz, M. T. Structural and Computational Insight into the Catalytic Mechanism of Limonene Epoxide Hydrolase Mutants in Stereoselective Transformations. *J. Am. Chem. Soc.* **2018**, *140* , 310–318.

(29) Li, G.; Zhang, H.; Sun, Z.; Liu, X.; Reetz, M. T. Multiparameter Optimization in Directed Evolution: Engineering Thermostability, Enantioselectivity, and Activity of an Epoxide Hydrolase. *ACS Catal.* **2016**, *6*, 3679–3687.

(30) Hopmann, K. H.; Hallberg, B. M.; Himo, F. Catalytic Mechanism of Limonene Epoxide Hydrolase, a Theoretical Study. *J. Am. Chem. Soc.* **2005**, *127*, 14339–14347.

(31) Lind, M. E. S.; Himo, F. Quantum Chemistry as a Tool in Asymmetric Biocatalysis: Limonene Epoxide Hydrolase Test Case. *Angew. Chemie - Int. Ed.* **2013**, *52*, 4563–4567.

(32) Hou, Q. Q.; Sheng, X.; Wang, J. H.; Liu, Y. J.; Liu, C. B. QM/MM Study of the Mechanism of Enzymatic Limonene 1,2-Epoxide Hydrolysis. *Biochim. Biophys. Acta - Proteins Proteomics* **2012**, *1824*, 263–268.

(33) Henzler-Wildman, K.; Kern, D. Dynamic Personalities of Proteins. *Nature* **2007**, *450*, 964–972.

(34) van der Kamp, M. W.; Prentice, E. J.; Kraakman, K. L.; Connolly, M.; Mulholland, A. J.; Arcus, V. L. Dynamical Origins of Heat Capacity Changes in Enzyme-Catalysed Reactions. *Nat. Commun.* **2018**, *9*, 1177.

(35) Goodey, N. M.; Benkovic, S. J. Allosteric Regulation and Catalysis Emerge via a Common Route. *Nat. Chem. Biol.* **2008**, *4*, 474–482.

(36) Frauenfelder, H.; Chen, G.; Berendzen, J.; Fenimore, P. W.; Jansson, H.; McMahon, B. H.; Stroe, I. R.; Swenson, J.; Young, R. D. A Unified Model of Protein Dynamics. *Proc. Natl. Acad.*

Sci. U. S. A. **2009**, *106*, 5129–5134.

(37) Bahar, I.; Lezon, T. R.; Yang, L.-W.; Eyal, E. Global Dynamics of Proteins: Bridging Between Structure and Function. *Annu. Rev. Biophys.* **2010**, *39*, 23–42.

(38) Smock, R. G.; Gierasch, L. M. Sending Signals Dynamically. *Science* **2009**, *324*, 198–203.

(39) Tsai, C.-J.; del Sol, A.; Nussinov, R. Protein Allostery, Signal Transmission and Dynamics: A Classification Scheme of Allosteric Mechanisms. *Mol. Biosyst.* **2009**, *5*, 207.

(40) Tokuriki, N.; Tawfik, D. S. Protein Dynamism and Evolvability. *Science* **2009**, *324*, 203–207.

(41) Csermely, P.; Palotai, R.; Nussinov, R. Induced Fit, Conformational Selection and Independent Dynamic Segments: An Extended View of Binding Events. *Trends Biochem. Sci.* **2010**, *35*, 539–546.

(42) Paladino, A.; Marchetti, F.; Rinaldi, S.; Colombo, G. Protein Design: From Computer Models to Artificial Intelligence. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2017**, *7*, 1–21.

(43) Serrano-Hervás, E.; Garcia-Borràs, M.; Osuna, S. Exploring the Origins of Selectivity in Soluble Epoxide Hydrolase from *Bacillus Megaterium*. *Org. Biomol. Chem.* **2017**, *15*, 8827–8835.

(44) Woods, C. J.; Malaisree, M.; Hannongbua, S.; Mulholland, A. J. A Water-Swap Reaction Coordinate for the Calculation of Absolute Protein–ligand Binding Free Energies. *J. Chem. Phys.* **2011**, *134*, 54114.

(45) Wijma, H. J.; Floor, R. J.; Bjelic, S.; Marrink, S. J.; Baker, D.; Janssen, D. B. Enantioselective Enzymes by Computational Design and In Silico Screening. *Angew. Chemie Int. Ed.* **2015**, *54* , 3726–3730.

(46) Johansson, P.; Unge, T.; Cronin, A.; Arand, M.; Bergfors, T.; Jones, T. A.; Mowbray, S. L. Structure of an Atypical Epoxide Hydrolase from Mycobacterium Tuberculosis Gives Insights into Its Function. *J. Mol. Biol.* **2005**, *351*, 1048–1056.

(47) Rinaldi, S.; Gori, A.; Annovazzi, C.; Ferrandi, E. E.; Monti, D.; Colombo, G. Unraveling Energy and Dynamics Determinants to Interpret Protein Functional Plasticity: The Limonene-1,2-Epoxide-Hydrolase Case Study. *J. Chem. Inf. Model.* **2017**, *57*, 717–725.

(48) Morra, G.; Neves, M. A. C.; Plescia, C. J.; Tsustsumi, S.; Neckers, L.; Verkhivker, G.; Altieri, D. C.; Colombo, G. Dynamics-Based Discovery of Allosteric Inhibitors: Selection of New Ligands

for the C-Terminal Domain of Hsp90. *J. Chem. Theory Comput.* **2010**, *6*, 2978–2989.

(49) Morra, G.; Verkhivker, G.; Colombo, G. Modeling Signal Propagation Mechanisms and Ligand-Based Conformational Dynamics of the Hsp90 Molecular Chaperone Full-Length Dimer. *PLoS Comput. Biol.* **2009**, *5*, e1000323.

(50) Morra, G.; Potestio, R.; Micheletti, C.; Colombo, G. Corresponding Functional Dynamics across the Hsp90 Chaperone Family: Insights from a Multiscale Analysis of MD Simulations. *PLoS Comput. Biol.* **2012**, *8*, e1002433.

(51) Rehn, A.; Moroni, E.; Zierer, B. K.; Tippel, F.; Morra, G.; John, C.; Richter, K.; Colombo, G.; Buchner, J. Allosteric Regulation Points Control the Conformational Dynamics of the Molecular Chaperone Hsp90. *J. Mol. Biol.* **2016**, *428*, 4559–4571.

(52) Garcia-Viloca, M.; Gao, J.; Karplus, M.; Truhlar, D. G. How Enzymes Work: Analysis by Modern Rate Theory and Computer Simulations. *Science.* **2004**, *303*, 186–195.

(53) Lodola, A.; Mor, M.; Zurek, J.; Tarzia, G.; Piomelli, D.; Harvey, J. N.; Mulholland, A. J. Conformational Effects in Enzyme Catalysis: Reaction via a High Energy Conformation in Fatty Acid Amide Hydrolase. *Biophys. J.* **2007**, *92* , L20–L22.

(54) Tsou, C.-L. Conformational Flexibility of Enzyme Active Sites. *Science .* **1993**, *262* , 380–382.

(55) Garcia-Viloca, M.; Gao, J.; Karplus, M.; Truhlar, D. G. A Perspective on Enzyme Catalysis. *Science.* **2004**, *301*, 1196–1202.

(56) Ranaghan, K. E.; Mulholland, A. J. Conformational Effects in Enzyme Catalysis: QM/MM Free Energy Calculation of the “NAC” Contribution in Chorismate Mutase. *Chem. Commun.* **2004**, *0* , 1238–1239.

(57) Claeyssens, F.; Harvey, J. N.; Manby, F. R.; Mata, R. A.; Mulholland, A. J.; Ranaghan, K. E.; Schütz, M.; Thiel, S.; Thiel, W.; Werner, H.-J. High-Accuracy Computation of Reaction Barriers in Enzymes. *Angew. Chemie* **2006**, *118*, 7010–7013.

(58) Himo, F. Quantum Chemical Modeling of Enzyme Active Sites and Reaction Mechanisms. *Theor. Chem. Acc.* **2006**, *116*, 232–240.

(59) Kästner, J.; Senn, H. M.; Thiel, S.; Otte, N.; Thiel, W. QM/MM Free-Energy Perturbation Compared to Thermodynamic Integration and Umbrella Sampling: Application to an Enzymatic Reaction. *J. Chem. Theory Comput.* **2006**, *2*, 452–461.

(60) Warshel, A. Computer Simulations of Enzyme Catalysis: Methods, Progress, and Insights. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 425–443.

(61) Acevedo, O.; Jorgensen, W. L. Advances in Quantum and Molecular Mechanical (QM/MM) Simulations for Organic and Enzymatic Reactions. *Acc. Chem. Res.* **2010**, *43*, 142–151.

(62) Friesner, R. A.; Guallar, V. Ab Initio Quantum Chemical and Mixed Quantum Mechanics/molecular Mechanics (QM/MM) Methods for Studying Enzymatic Catalysis. *Annu. Rev. Phys. Chem.* **2005**, *56*, 389–427.

(63) Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chemie Int. Ed.* **2009**, *48*, 1198–1229.

(64) van der Kamp, M. W.; Mulholland, A. J. Combined Quantum Mechanics/Molecular Mechanics (QM/MM) Methods in Computational Enzymology. *Biochemistry* **2013**, *52*, 2708–2728.

(65) Field, M. J.; Bash, P. A.; Karplus, M. A Combined Quantum Mechanical and Molecular Mechanical Potential for Molecular Dynamics Simulations. *J. Comput. Chem.* **1990**, *11*, 700–733.

(66) Warshel, A.; Levitt, M. Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme. *J. Mol. Biol.* **1976**, *103* , 227–249.

(67) Cui, Q.; Elstner, M.; Kaxiras, E.; Frauenheim, T.; Karplus, M. A QM/MM Implementation of the Self-Consistent Charge Density Functional Tight Binding (SCC-DFTB) Method. *J. Phys. Chem. B* **2000**, *105*, 569–585.

(68) Lodola, A.; Mor, M.; Sirirak, J.; Mulholland, A. J. Insights into the Mechanism and Inhibition of Fatty Acid Amide Hydrolase from Quantum Mechanics/molecular Mechanics (QM/MM) Modelling. *Biochem. Soc. Trans.* **2009**, *37*, 363–367.

(69) van der Kamp, M. W.; Perruccio, F.; Mulholland, A. J. Ab Initio QM/MM Modelling of Acetyl-CoA Deprotonation in the Enzyme Citrate Synthase. *J. Mol. Graph. Model.* **2007**, *26*, 676–690.

(70) Bowman, A. L.; Grant, I. M.; Mulholland, A. J. QM/MM Simulations Predict a Covalent Intermediate in the Hen Egg White Lysozyme Reaction with Its Natural Substrate. *Chem. Commun.* **2008**, *0*, 4425.

(71) Dudev, T.; Lim, C. Ring Strain Energies from Ab Initio Calculations. *J. Am. Chem. Soc.* **1998**,

120, 4450–4458.

(72) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D.E.; Francis, P.; Shenkin, P.S. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **2004**, *47*, 1739–1749.

(73) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides. *J. Phys. Chem. B* **2001**, *105*, 6474–6487.

(74) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins Struct. Funct. Bioinforma.* **2006**, *65*, 712–725.

(75) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.

(76) Loncharich, R. J.; Brooks, B. R.; Pastor, R. W. Langevin Dynamics of Peptides: The Frictional Dependence of Isomerization Rates of N-Acetylalanyl-N'-Methylamide. *Biopolymers* **1992**, *32* , 523–535.

Table of Contents